

1/25

1 MGS LFQEAEP QAGTEQNKPT LASRFQQT LG DLLARLGSRG HVYVIHCLNP
51 TPGKIPGLLD VGHVAEQLRQ AGILEIIGTR STHFPVRVSF QVFLARFHAL
101 GSGRQKAASD QERCGAILSE VLGAESPLYH LGVTQVLLQE QGWQQLEQLW
151 AQRRSQALLT LHRGLRACIT RQRLRLLPRM QARVRGLQAR KRYLQRRSAL
201 GQLNTILLVA RPLLRRRQKL RCAPGPHSGE PWGKVSNDL GRLEIPAQLA
251 TLLERAEGHQ ALLTG SITES LPPEV PARPS LTLPPDIDQF PFSSFVSTSF
301 QKPFLPRPGQ PLDEPLTRLD GENPQQA LEI NRVMLRLLGE GSLQSWQEQT
351 MGTFLVQQAQ RRPGLRDELF SQLVAQLWRN PDEQQNQRGW ALMVILLSSF
401 APTPALEKPL LKFVSDQAPS GMAALCQ HKL LGALEQTPLA PMASRSHPPT
451 QLEWKAGLRR GRMALDVFTF NEESYS AEVE SWTTGEQFAG WILQSRGLEA
501 PPRGWSVSLH SGDAWRDLPG CDFVLDLIGQ TEDLGDPAGP HNY PITPLGL
551 AESIPPAPGV QAPSLPPGLP PGPAPILASS RPPGEASKPE NLDGFVDHLE
601 EPALAPGFSD LEQGWALSRR MKGGG SVGPT QQGYPMVYPG MVQAPSYQPA
651 MIPAPMPVMP AMGAVPTMPA MMVPPQPQPL VPSLDSRQLA LQQQNF INQQ
701 AMILAQQMTT QAMSLSLEQQ NQRHQHQAQT SGATSQPPPS TTAPKAKKPP
751 APQEK PESNL EPSGVGLRED TPEEAESKPQ RPKSFQQKRD YFQKMGQDPI
801 RVKTVKPPAK VQIPQEEMEE TEE EDETAE LSPPPPPPPV VKKPLKASRP
851 KAVKEDEAEP AQEEVPTQGE DPPVHSSNSA PQHPKPSRVP PVQSSNSAPP
901 RPQPSREIRN IIRMYQSRPG PVAVPVQPTR PIKTFQKKND PKDEALAKLG
951 INGVHLPLST SPNQGKSSPP AVVPRPKARP RLEPSLSIQE KQG PLRDLFG
1001 PCSPNPPTAP APPPPPALPP PLSGEPKTPS VESHALTEPM EDKNISTKLL
1051 VPSGSVCFSY ANAPWKFLR KEVFYPRENF SHPYCLSLLC QQILRDTFTE
1101 SCTRISQDER HKMKGLL GDL EVSLETLDIV EDSIKKRIVV AARDNWANYE

FIG. 1A.

2/25

+

1151 SRIFPVSGES GSDVQLLGVS HRGLRLLKVT QSPSFHLDQL KTLCSYSYAE
1201 VLTVQCRGRS TLELSLKNEQ LILHTAWARA IKAMVDLFLS ELRKDSGYVI
1251 ALRSYITDDN SLLSFHRGDL IRLLPVTALE PGWQFGSAGG RSGLFPDDVV
1301 QPAAAPDLSF SLGKRNSWQR KSKLGPAQEV RKTEEVK*

FIG. 1B.

1151 SRIFPVSGES GSDVQLLGVS HRGLRLLKVT QSPSFHLDQL KTLCSYSYAE
1201 VLTVQCRGRS TLELSLKNEQ LILHTAWARA IKAMVDLFLS ELRKDSGYVI
1251 ALRSYITDDN SLLSFHRGDL IRLLPVTALE PGWQFGSAGG RSGLFPDDVV
1301 QPAAAPDLSF SLGKRNSWQR KSKLGPAQEV RKTEEVK*

+

3/25

1 CGCTGGGACT GTCACCTACC AGGTGCACAA GTTCATAAAC AGAAACAGGG
51 GCCACCTGGA CCCCCTGTG CTGGAGATGC TCAGGCAGAG CCAGCTGCAG
101 GTGACCTAGC CTTCTTTCA GTCATGGGC AGCCTGTTCC AAGAAGCAGA
151 GCCCCAGGCT GGGACTGAGC AAAACAAACC CACATTGGCC TCTCGATTCC
201 AGCAGACCCT GGGTGACTTG CTAGCTCGGC TAGGCAGCAG GGGCCATGTC
251 TACGTCATCC ACTGTCTCAA TCCCACCCCT GGAAAGATCC CAGGCCTCTT
301 GGACGTGGGG CATGTGGCAG AGCAGCTGCG TCAGGCTGGC ATCCTGGAGA
351 TCATAGGCAC CCGGAGTACC CACTTCCCCG TCGAGTGTC CTTCCAAGTC
401 TTTCTGGCAA GGTTCATGC CCTGGGGTCA GGGAGACAGA AAGCTGCCTC
451 TGACCAGGAG AGGTGTGGTG CCATCCTCAG TGAAGTGCTG GGGGCAGAGT
501 CACCGCTGTA TCATCTTGA GTCACCCAGG TCCTGCTGCA GGAACAGGGC
551 TGGCAGCAGC TAGAACAGCT GTGGGCTCAG CGGCGCTCAC AGGCCCTGCT
601 CACTCTGCAC CGTGGCCTCC GAGCCTGTAT CACCCGGCAG CGCCTCCGTC
651 TCCTGCCCCG GATGCAGGCT CGTGTGCGTG GGCTCCAGGC CAGGAAGCGA
701 TATCTCCAGC GGAGGTCAGC TCTGGGACAG CTGAACACCA TTCTCCTAGT
751 GGCCCGGCCC CTGCTCCGGA GACGACAGAA GCTACGGTGT GCCCCTGGCC
801 CGCACAGCGG GGAGCCCTGG GGGAAAGTGT CAAATATGGA CCTGGGTCGC
851 TTAGAGATCC CCGCCCAGCT GGCTACTCTG CTGGAGAGGG CGGAAGGCCA
901 CCAGGCCTTG CTGACGGGGA GCATCACAGA GTCCCTGCCA CCTGAGGTCC
951 CCGCCCGGCC CAGCCTGACT CTCCCTCCAG ACATTGACCA GTTTCCTTC
1001 TCCAGTTTTG TATCCACCAG CTTTCAGAAG CCATTTCTGC CTCGACCAGG
1051 GCAGCCACTG GACGAGCCCC TGACGCGGTT AGATGGCGAG AACCCCTCAGC

FIG. 2A.

4/25

+

1101 AGGCTCTGGA GATCAACAGG GTGATGCTGC GGCTCCTGGG GGAAGGATCT
1151 CTGCAGTCCT GGCAAGAGCA GACCATGGGC ACGTTCCTCG TGCAGCAGGC
1201 CCAGCGACGG CCGGGACTCC GAGATGAGCT CTTAGCCAG CTGGTGGCCC
1251 AGCTGTGGCG CAACCCAGAT GAGCAACAGA ATCAGCGTGG CTGGGCCCTA
1301 ATGGTGATCC TGCTCAGCTC CTTTGCTCCC ACACCTGCCC TGGAGAAGCC
1351 ACTGCTCAAA TTTGTATCTG ACCAGGCTCC CAGTGGCATG GCAGCCCTGT
1401 GCCAGCACAA GCTGTTAGGT GCCCTGGAGC AGACACCGCT GGCTCCCATG
1451 GCTTCGAGGT CCCACCCACC CACACAACCTT GAGTGGAAGG CTGGTTTACG
1501 TCGGGGCCGC ATGGCGCTGG ATGTGTTTAC ATTCAACGAG GAAAGCTACT
1551 CCGCGGAAGT GGAATCCTGG ACCACGGGAG AGCAGTTTGC AGGGTGGATC
1601 CTACAGAGCA GAGGCCTGGA GCGCCCCCT CGTGGCTGGT CTGTGTCACT
1651 GCATTCTGGG GATGCTTGGC GTGACTTGCC TGGCTGTGAC TTTGTGTTGG
1701 ACCTAATAGG CCAGACTGAG GACTTGGGAG ACCCAGCTGG TCCCCACAAC
1751 TACCCCATCA CTCCTCTTGG TTTAGCTGAG AGCATCCCTC CAGCCCCTGG
1801 TGTCCAGGCT CCTTCCCTGC CCCCAGGACT CCCTCCAGGT CCAGCCCCAA
1851 TACTGGCCAG CAGCCGCCCT CCGGGCGAGG CCAGTAAGCC TGAGAACCTG
1901 GATGGTTTCG TGGACCACCT CTTTGAACCA GCGCTCGCTC CGGGTTTCAG
1951 TGATCTGGAA CAAGGCTGGG CCCTGAGCAG ACGCATGAAG GGAGGGGGCT
2001 CTGTTGGGCC CACCCAGCAG GGCTACCCCA TGGTGTACCC AGGTATGGTG
2051 CAGGCACCTA GCTACCAGCC AGCTATGATA CCCGCACCGA TGCCCGTCAT
2101 GCCAGCCATG GGCGCAGTCC CAACCATGCC AGCCATGATG GTGCCACCCC
2151 AGCCACAGCC TCTGGTGCCC AGTTTGGACT CAAGGCAGCT GGC ACTACAG
2201 CAGCAAACT TCATCAACCA GCAGGCGATG ATTCTGGCGC AGCAGATGAC
2251 CACCCAGGCC ATGAGCCTGT CCCTGGAGCA GCAGAATCAG AGACACCAGC

FIG. 2B.

+

5/25

2301 ACCAAGCTCA GACCTCTGGG GCCACCTCCC AGCCTCCACC CTCAACCACT
2351 GCTCCCAAGG CCAAGAAGCC TCCTGCCCCC CAAGAGAAGC CAGAGAGTAA
2401 CCTAGAGCCT TCGGGTGTTG GCTTGAGAGA GGACACCCCA GAGGAAGCTG
2451 AAAGCAAGCC TCAGCGCCCC AAGAGCTTCC AACAGAAACG GGACTATTTC
2501 CAGAAGATGG GGCAAGATCC GATCAGAGTG AAGACGGTGA AACCTCCAGC
2551 CAAGGTTTCA ATCCCCAAG AGGAGATGGA GGAGACGGAG GAGGAGGAGG
2601 ATGAGACCGC CGAGTTGTCC CCTCCTCCTC CCCCTCCCCC GGTGTGAAG
2651 AAGCCGCTGA AGGCAAGCAG GCCCAAAGCC GTAAAGGAAG ATGAGGCAGA
2701 GCCCGCCCAG GAGGAAGTAC CGACCCAGGG CGAGGATCCC CCGGTGCACA
2751 GCTCCAATC CGCACCTCAG CACCCCAAAC CCAGCAGGGT ACCCCCAGTG
2801 CAGAGCTCCA ACTCCGCACC TCCACGCCCG CAACCCAGCA GGGAAATCCG
2851 AAACATCATC CGAATGTACC AGAGCCGTCC AGGGCCTGTG GCTGTGCCCC
2901 TACAACCCAC CAGGCCCATC AAAACTTTTC AGAAGAAAAA TGACCCTAAG
2951 GATGAGGCTT TGGCTAAGTT AGGGATAAAT GCGTCCACT TGCCCCTATC
3001 GACATCGCCT AACCAAGGGA AGAGCTCTCC ACCGGCTGTA GTTCCTCGAC
3051 CTAAGGCTCG ACCTCGTCTT GAGCCTTCCC TATCCATCCA GGAAAAGCAG
3101 GGACCCCTTC GGGACTTGTT TGGCCCATGT AGTCCAAACC CACCTACAGC
3151 TCCAGCACCC CCGCCTCCAC CAGCACTCCC ACCGCCTCTG TCTGGGGAGC
3201 CCAAGACCCC TTCAGTGGAG TCTCATGCCT TGACAGAGCC CATGGAGGAC
3251 AAGAACATCT CCACAAAGCT CCTTGTGCCC TCTGGAAGTG TGTGCTTCTC
3301 CTATGCCAAT GCACCCTGGA AGTTGTTCTT ACGCAAGGAG GTGTTCTACC
3351 CCCGGGAGAA CTTCAATCAT CCATACTGCC TCAGTCTCCT CTGCCAGCAG
3401 ATCCTGCGGG ACACCTTCAC AGAGTCCTGC ACCCGGATCT CACAGGATGA

FIG. 2C.

6/25

3451 GCGGCACAAA ATGAAAGGCC TTCTGGGAGA CTTGGAGGTG AGTCTGGAGA
3501 CCCTTGACAT TGTGAAGAC AGCATCAAAA AACGCATCGT GGTCGCTGCT
3551 CGGGACAAC TGGCCAATTA CTTCTCCCGC ATCTTCCCAG TCTCGGGTGA
3601 GAGTGGCAGC GATGTACAGC TGCTGGGTGT GTCTACCGG GGAAGTGGGC
3651 TGCTGAAGGT GACCCAAAGC CCGAGCTTCC ACCTGGACCA GCTGAAGACA
3701 CTCTGTTTCT ACAGCTATGC TGAAGTCCTG ACCGTGCAGT GCAGGGGCAG
3751 ATCCACCCTG GAGCTGTCCT TGAAGAATGA GCAGCTGATA CTGCACACAG
3801 CCTGGGCGAG GGCCATCAAG GCCATGGTGG ATCTATTTCT GAGTGAAGTC
3851 AGGAAGGACT CCGGCTATGT CATCGCCCTG CGCAGCTACA TCACCGATGA
3901 CAATAGCCTC CTCAGTTTCC ACCGTGGGGA CCTCATTAGG TTAGTGCCAG
3951 TGACCGCTCT GGAACCAGGC TGGCAGTTCG GTTCTGCCGG GGGCCGCTCC
4001 GGACTCTTTC CCGATGACGT GGTGCAGCCA GCTGCTGCCC CCGACCTCTC
4051 CTTTTCCCTG GGAAAGAGAA ACAGCTGGCA ACGCAAGAGT AAGCTGGGGC
4101 CAGCTCAGGA GGTGAGGAAG ACAGAAGAGG TGAAGTGATA CAGGCCTAAC
4151 TTGGAGACTG AGAAGGAAAG AGCAGGGTTG CTTCCGGGTGT TGTCCACTTC
4201 CTGTCCTGGT GGCCAGGGCT CAATGTGTTC CTGTCCTTTA CCATCTCCTG
4251 ACTTTTTGCC ATTTGTGAGA CTGTAAGTCA CACCCTCTAA CTCTGGTACT
4301 TAGTTCAGTG TCTCCATAGA GGATGCTTAA TAAATAACCT TGGTTTTCTC
4351 GGTTTCTGGT GTCACCTCTC TTGGGTCTAA TGGGTATGGG GACCAGGGCC
4401 TGAGAGTGAG TATTGGGCCT CTGGGCTAGA TGGTGGGTAC TGGGGTGGTA
4451 CCAAATTTCC TGTGCTCCCA GCGCCCCACC CATCCCAGGA AACAGAACC
4501 CAGTGAAGAC TCGGAGGCCA CCTCCTTTAC AACCTACAGC TCTTTGTCTG
4551 CCGACCCCCA CAACTACACC ATGCAGGAAT TTGCCCTGCG CTATTTCCGG
4601 AAGCCTCATA CCTGGCTGAC CCAGATGAGT AGAGACACCA AAGAGAAAGC

FIG. 2D.

+

7/25

4651 TGCCATCAAC CTGATCCAGT AACTAAGGA CCCCATCCAG GAATCCCTTA
4701 CCAGCTTCTG CAATGGGGAC ACAAACAGTA AAGCTGTGGC TGGCTTCAAG
4751 GCTCTGATGC AGTTTATGGG GGACCAGCCT AAGCCCCGGG GCAAGGACGA
4801 GCTGAGTCTG CTCTATGAGC TGCTGAAGCT GTGCCAAGAT GACCTTAGGG
4851 ACGAGATGTA CTGCCAGGTC ATCAAGCAAG TCACAGGACA CCCCCAGCCA
4901 AAGCACTGTG CTCTGGGCTG GAGCGTCCTC AGCCTCTTCA CAGGCTTCTT
4951 TGCACCATCG ACCACGCTGA TGCCCTATGT GACCAAGTTC CTGCAGGATT
5001 CCAGCCCCAG TGAAGAGTTG GCCAGGAGGA GCCAGGAGAA CCTCCAGCGC
5051 ACAGTTAAAT ATGGGGGACG CCAGCAGCTG CCGTTACCTG GTGAAATGAA
5101 TGCTTTTCTG AAAGGGCAAG CAGTTCGTTT GCTTCTAATT CACCTGCCTG
5151 GGGGTGTGGA CTACAGGACG AATTCACAGA CATTACAGT GGCAGGGGAA
5201 GTGCTAGAGG AGCTGTGTGG ACAGATGGGC ATCACAGACT TGAAGAAGT
5251 GCAGGAATTT GCCCTCTTTC TCATCAAAGG AGAAGGTGAG CTGGTTCGGC
5301 CGCTGTCACC CCATGAGTAC ATCAACAATG TGGTGACGGA CCAGGACATG
5351 AGCCTTCACA GCCGACGGCT TGGTTGGGAG ACTCCACTGC ATTTTGATCA
5401 CTCCACCTAC ACGGAAACCC ACTATGGCCA GGTGCTTCGG GACTACCTGC
5451 AAGGGAAGCT GATAGTCAGC ACCCAGGCAG AGGCTCTACT TGCCCAGCTT
5501 GCTGCCTTCC AACACTTCGA CAAAACCGGA ACTTCTAGTC CTCCATCAGA
5551 GCAAGAGCTG CTGTCTTATA TTCCCAAGCC ACTGCAATGG CAGGTGAACA
5601 CAGCCAACAT AAAGAGCTTG GTGACCCAGG AGCTGAGGCA GATGCAAGGG
5651 TACAGCAAGC AGAGAGCACA GATTGGCTTT ATAGAGAGCA CAGCGCAGCT
5701 GCCCCTCTTT GGCTACACTG TGTACGTAGT GCTGAGAGTG AGTAAGCTGG
5751 CCCTCCCTGG ACCAGGCCTC CTGGGGCTGA ACCGTCAGCA CCTGGTCCTC

FIG. 2E.

+

8/25

5801 ATGGACCCCA GCTCTCAGGA ACTCTGCTGC TCTGTCATGC TAAAAGACCT
5851 GAAGCAGTTC CACCTGCTGA GCCCACTGCA GGAGGACGGG CCCCTGGCC
5901 TAGAACTCAA CTATGGCTCT GTTGACAACC CCCAGACCAT CTGGTTGGAG
5951 TTGCCACAGG CCCAGGAGCT GCAGCACACC ATCATCTTCC TGCTGGGCAG
6001 CATGTCCACT CAGTGGCCAG GTCTCCTCTG AGGAGTGGAG ATAAGGCAGC
6051 GGTCTCTCAC TGGGCAGTCT GCCTTAGTCC TGCTCTGAAT CCGCTGCACA
6101 ACCCCCCACC CCACGTGGAG GCCAAAAGGC AAAGTTGTGT CACCTGGGAG
6151 AATAGGCAGA CACATCCCCT CTGGGGTGGA CTGCAACAGG AGTTGGGGCA
6201 TTTGCTGGCT AGCCCCAGGG AAAATGCCCA CCCAGCTCGA AAGCGGCACA
6251 AGTAAACAC CCAAGGAAAA AAAAAAAAAA AAAAAAAAAA AAA

FIG. 2F.

9/25

+

1 CGCTGGGACT GTCACCTACC AGGTGCACAA GTTCATAAAC AGAAACAGGG
51 GCCACCTGGA CCCCCTGTG CTGGAGATGC TCAGGCAGAG CCAGCTGCAG
101 GTGACCTAGC CTTCTTTCA GTCATGGGC AGCCTGTTCC AAGAAGCAGA
151 GCCCCAGGCT GGGACTGAGC AAAACAAACC CACATTGGCC TCTCGATTCC
201 AGCAGACCCT GGGTGACTTG CTAGCTCGGC TAGGCAGCAG GGGCCATGTC
251 TACGTCATCC ACTGTCTCAA TCCCACCCCT GGAAAGATCC CAGGCCTCTT
301 GGACGTGGGG CATGTGGCAG AGCAGCTGCG TCAGGCTGGC ATCCTGGAGA
351 TCATAGGCAC CCGGAGTACC CACTTCCCCG TCGAGTGTC CTTCCAAGTC
401 TTTCTGGCAA GGTTCATGC CCTGGGGTCA GGGAGACAGA AAGCTGCCTC
451 TGACCAGGAG AGGTGTGGTG CCATCCTCAG TGAAGTGCTG GGGGCAGAGT
501 CACCGCTGTA TCATCTTGGA GTCACCCAGG TCCTGCTGCA GGAACAGGGC
551 TGGCAGCAGC TAGAACAGCT GTGGGCTCAG CGGCGCTCAC AGGCCCTGCT
601 CACTCTGCAC CGTGGCCTCC GAGCCTGTAT CACCCGGCAG CGCCTCCGTC
651 TCCTGCCCCG GATGCAGGCT CGTGTGCGTG GGCTCCAGGC CAGGAAGCGA
701 TATCTCCAGC GGAGGTCAGC TCTGGGACAG CTGAACACCA TTCTCCTAGT
751 GGCCCGGCCC CTGCTCCGGA GACGACAGAA GCTACGGTGT GCCCCTGGCC
801 CGCACAGCGG GGAGCCCTGG GGGAAAGTGT CAAATATGGA CCTGGGTCGC
851 TTAGAGATCC CCGCCAGCT GGCTACTCTG CTGGAGAGGG CGGAAGGCCA
901 CCAGGCCTTG CTGACGGGGA GCATCACAGA GTCCCTGCCA CCTGAGGTCC
951 CCGCCCGGCC CAGCCTGACT CTCCCTCCAG ACATTGACCA GTTTCCTTC
1001 TCCAGTTTTG TATCCACCAG CTTTCAGAAG CCATTTCTGC CTCGACCAGG
1051 GCAGCCACTG GACGAGCCCC TGACGCGGTT AGATGGCGAG AACCTCAGC

FIG. 3A.

+

10/25

1101 AGGCTCTGGA GATCAACAGG GTGATGCTGC GGCTCCTGGG GGAAGGATCT
1151 CTGCAGTCCT GGCAAGAGCA GACCATGGGC ACGTTCCTCG TGCAGCAGGC
1201 CCAGCGACGG CCGGGACTCC GAGATGAGCT CTTCAGCCAG CTGGTGGCCC
1251 AGCTGTGGCG CAACCCAGAT GAGCAACAGA ATCAGCGTGG CTGGGCCCTA
1301 ATGGTGATCC TGCTCAGCTC CTTTGCTCCC ACACCTGCCC TGGAGAAGCC
1351 ACTGCTCAAA TTTGTATCTG ACCAGGCTCC CAGTGGCATG GCAGCCCTGT
1401 GCCAGCACAA GCTGTTAGGT GCCCTGGAGC AGACACCGCT GGCTCCCATG
1451 GCTTCGAGGT CCCACCCACC CACACAACCTT GAGTGAAGG CTGGTTTACG
1501 TCGGGGCCGC ATGGCGCTGG ATGTGTTTAC ATTCAACGAG GAAAGCTACT
1551 CCGCGGAAGT GGAATCCTGG ACCACGGGAG AGCAGTTTGC AGGGTGGATC
1601 CTACAGAGCA GAGGCCTGGA GCGCCCCCT CGTGGCTGGT CTGTGTCACT
1651 GCATTCTGGG GATGCTTGGC GTGACTTGCC TGGCTGTGAC TTTGTGTTGG
1701 ACCTAATAGG CCAGACTGAG GACTTGGGAG ACCCAGCTGG TCCCCACAAC
1751 TACCCCATCA CTCCTCTTGG TTTAGCTGAG AGCATCCCTC CAGCCCCTGG
1801 TGTCCAGGCT CCTTCCCTGC CCCAGGACT CCCTCCAGGT CCAGCCCCAA
1851 TACTGGCCAG CAGCCGCCCT CCGGGCGAGG CCAGTAAGCC TGAGAACCTG
1901 GATGGTTTCG TGGACCACCT CTTTGAACCA GCGCTCGCTC CGGGTTTCA
1951 TGATCTGGAA CAAGGCTGGG CCCTGAGCAG ACGCATGAAG GGAGGGGGCT
2001 CTGTTGGGCC CACCCAGCAG GGCTACCCCA TGGTGTACCC AGGTATGGTG
2051 CAGGCACCTA GCTACCAGCC AGCTATGATA CCCGCACCGA TGCCCGTCAT
2101 GCCAGCCATG GGCGCAGTCC CAACCATGCC AGCCATGATG GTGCCACCCC
2151 AGCCACAGCC TCTGGTGCCC AGTTTGGACT CAAGGCAGCT GGCCTACAG
2201 CAGCAAACT TCATCAACCA GCAGGCGATG ATTCTGGCGC AGCAGATGAC
2251 CACCCAGGCC ATGAGCCTGT CCCTGGAGCA GCAGAATCAG AGACACCAGC

FIG. 3B.

11/25

+

2301 ACCAAGCTCA GACCTCTGGG GCCACCTCCC AGCCTCCACC CTCAACCACT
2351 GCTCCCAAGG CCAAGAAGCC TCCTGCCCCC CAAGAGAAGC CAGAGAGTAA
2401 CCTAGAGCCT TCGGGTGTG GCTTGAGAGA GGACACCCCA GAGGAAGCTG
2451 AAAGCAAGCC TCAGCGCCCC AAGAGCTTCC AACAGAAACG GGACTATTTC
2501 CAGAAGATGG GGCAAGATCC GATCAGAGTG AAGACGGTGA AACCTCCAGC
2551 CAAGGTTCAG ATCCCCCAAG AGGAGATGGA GGAGACGGAG GAGGAGGAGG
2601 ATGAGACCGC CGAGTTGTCC CCTCCTCCTC CCCCTCCCCC GGTGTGAAG
2651 AAGCCGCTGA AGGCAAGCAG GCCCAAAGCC GTAAAGGAAG ATGAGGCAGA
2701 GCCCGCCCAG GAGGAAGTAC CGACCCAGGG CGAGGATCCC CCGGTGCACA
2751 GCTCCAACTC CGCACCTCAG CACCCCAAAC CCAGCAGGGT ACCCCCAGTG
2801 CAGAGCTCCA ACTCCGCACC TCCACGCCCG CAACCCAGCA GGGAAATCCG
2851 AAACATCATC CGAATGTACC AGAGCCGTCC AGGGCCTGTG GCTGTGCCCC
2901 TACAACCCAC CAGGCCCATC AAAACTTTTC AGAAGAAAAA TGACCCTAAG
2951 GATGAGGCTT TGGCTAAGTT AGGGATAAAT GCGTCCACT TGCCCCTATC
3001 GACATCGCCT AACCAAGGGA AGAGCTCTCC ACCGGCTGTA GTTCCTCGAC
3051 CTAAGGCTCG ACCTCGTCTT GAGCCTTCCC TATCCATCCA GGAAAAGCAG
3101 GGACCCCTTC GGGACTTGTT TGGCCCATGT AGTCCAAACC CACCTACAGC
3151 TCCAGCACCC CCGCCTCCAC CAGCACTCCC ACCGCCTCTG TCTGGGGAGC
3201 CCAAGACCCC TTCAGTGGAG TCTCATGCCT TGACAGAGCC CATGGAGGAC
3251 AAGAACATCT CCACAAAGCT CCTTGTGCCC TCTGGAAGTG TGTGCTTCTC
3301 CTATGCCAAT GCACCCTGGA AGTTGTTCTT ACGCAAGGAG GTGTTCTACC
3351 CCCGGGAGAA CTTCAATCAT CCATACTGCC TCAGTCTCCT CTGCCAGCAG
3401 ATCCTGCGGG ACACCTTCAC AGAGTCCTGC ACCCGGATCT CACAGGATGA

FIG. 3C.

+

12/25

3451 GCGGCACAAA ATGAAAGGCC TTCTGGGAGA CTTGGAGGTG AGTCTGGAGA
3501 CCCTTGACAT TGTTGAAGAC AGCATCAAAA AACGCATCGT GGTCTGCTGCT
3551 CGGGACAACCT GGGCCAATTA CTTCTCCCGC ATCTTCCCAG TCTCGGGTGA
3601 GAGTGGCAGC GATGTACAGC TGCTGGGTGT GTCTCACCGG GGAAGTGGGC
3651 TGCTGAAGGT GACCCAAAGC CCGAGCTTCC ACCTGGACCA GCTGAAGACA
3701 CTCTGTTTCT ACAGCTATGC TGAAGTCCTG ACCGTGCAGT GCAGGGGCAG
3751 ATCCACCCTG GAGCTGTCCT TGAAGAATGA GCAGCTGATA CTGCACACAG
3801 CCTGGGCGAG GGCCATCAAG GCCATGGTGG ATCTATTTCT GAGTGAAGTC
3851 AGGAAGGACT CCGGCTATGT CATCGCCCTG CGCAGCTACA TCACCGATGA
3901 CAATAGCCTC CTCAGTTTCC ACCGTGGGGA CCTCATTAGG TTAGTGCCAG
3951 TGACCGCTCT GGAACCAGGC TGGCAGTTCG GTTCTGCCGG GGGCCGCTCC
4001 GGAAGCTTTT CCGATGACGT GGTGCAGCCA GCTGCTGCCC CCGACCTCTC
4051 CTTTTCCCTG GGAAAGAGAA ACAGCTGGCA ACGCAAGAGT AAGCTGGGGC
4101 CAGCTCAGGA GGTGAGGAAG ACAGAAGAGG TGAAGTGATA CAGGCCTAAC
4151 TTGGAGACTG AGAAGGAAAG AGCAGGGTTG CTTGGGGTGT TGTCCACTTC
4201 CTGTCCTGGT GGCCAGGGCT CAATGTGTTC CTGTCCTTTA CCATCTCCTG
4251 ACTTTTTGCC ATTTGTGAGA CTGTAAGTCA CACCCTCTAA CTCTGGTACT
4301 TAGTTCAGTG TCTCCATAGA GGATGCTTAA TAAATAACCT TGGTTTTCTC
4351 GGAAAAAAAA AAAAAAAAAA AAAAA

FIG. 3D.

13/25

MYQSRPGVPVPVQPSRPPKAFLRKIDPKDEALAKLINGAHSSPPMLSPSPGKGPPPAVAPRPKA
PLQLGPSSSIKEKQGPIILDLFQKLPPIAHTPPPPAPPPLPLPEDPGTISAERRCLTQPVEDQGVST
QLLAPSGSVCFSYTGTPWKLFRLRKEVFYPRENESHPHYVLRLLCEQILRDTFSESCIRISQNERKKM
KDLLGGLVDLDSLTTTSDSVKKRIVVAARDNWANYFSRFFPVSGESGSDVQLLAVSHRGLRLKV
TQGPGLRPDQLKILCSYSFAEVLGVECRGGSTLELSLKSEQLVHTARARAIEALVELFLNELKKD
SGYVIALRSYITDNCSLLSFHRGDLIKLLPVATLEPGWQFGSAGRSGLFPADIVQPAAAPDFSFS
KEQRSQWHKGGQLSNGEPGLARWDRASEVRKMGEQAEARPA

FIG. 4.

FIG. 5A.

15/25

+

TCCCACGTGGTACTTAGTTCAAGGCTGCCCCAGCAGATGCTTAATAAACAGCTCTTCACTTTTCCTG
GCTTCTGGTCTTGCTCCTTTGGTGTCTGGCTGGGGAGGGATGGGGCTGGGGCAGGACCCCTGGGAC
AGGGCACTGGACACTCAGGTGGCACCAGGTTTCTTGTGATCCCAGCGCCCTGCCCACCCTTGGAGC
CAGGCACACAGTGACGACTCGGAGGCCACCAGCCTGTCTCTGTGGCCTATGCCTTTCTGCCCCGAC
TCCCACAGCTACACCATGCAGGAATTCGCCCCGGCGTTACTTCCGGAGGTCCCAGGCCTTGCTGGGC
CAGACTGATGGAGGTGCCGCAGGAAAGGACACGGACAGCCTGGTGCAGTACACCAAGGCTCCCATC
CAGGAGTCGCTCCTCAGCCTCAGTGATGATGTGAGCAAGCTGGCTGTAGCCAGCTTCCTGGCCCCCT
GATGCGGTTTATGGGTGACCAGTCCAAGCCCCGGGGCAAGGATGAGATGGATCTGCTCTATGAACT
GCTGAAGCTGTGCCAGCAGGAGAAGCTGAGGGATGAGATTTACTGCCAGGTTATCAAGCAGGTCAC
AGGACACCCCCGGCCGGAACACTGCACTCGAGGCTGGAGCTTCCTCAGCCTTCTCACAGGCTTCTT
CCCCCGTTCGACCAGGCTGATGCCCTACCTGACCAAGTTTCTGCAGGATTCAGGCCCCAGCCAAGA
GCTGGCCCCGGAGCAGCCAGGAGCACCTCCAGCGCACAGTCAAATATGGGGGGCGCCGGCGGATGCC
CCCACCGGGTGAAATGAAGGCTTTCCTGAAAGGACAAGCGATTTCGCTGCTTCTTATTCACCTGCC
GGGGGGTGTGGATTATAGGACGAATATCCAGACTTTCACAGTAGCAGCAGAAGTGCAGGAGGAGCT
GTGCCGGCAAATGGGTATCACGGAGCCTCAGGAAGTGCAGGAATTCGCCCTCTTCCTCATCAAAGA
GAAGAGCCAGCTGGTTCGGGCCCTGCAGCCCGCCGAATACCTCAACAGCGTGGTAGTGACCAGGA
CGTGAGCCTGCACAGCCGGCGGCTCCACTGGGAGACCCCACTGCACTTCGATAACTCCACCTACAT
CAGCACCCACTACAGCCAGGTGCTGTGGGACTACCTTCAGGGGAAGCTGCCAGTCAGCGCCAAGGC
AGACGCGCAGCTCGCCAGGCTGGCCGCCCTGCAGCACCTCAGCAAGGCCAACAGGAATACCCCTC
AGGGCAGGACCTGCTAGCTTACGTGCCAAAGCAGCTGCAACGGCAGGTGAACACGGCCTCCATCAA
GAACCTGATGGGTGAGGAGCTGAGACGGCTGGAAGGACACAGCCCCCAGGAAGCACAGATCAGCTT
CATTGAGGCCATGAGCCAGCTGCCCTCTTCGGCTACACCGTCTATGGGGTGCTGCGAGTGAGCAT
GCAGGCCCTGTCCGGACCCACTCTCCTGGGGCTCAACCGCCAGCATCTCATCCTCATGGACCCAG
CTCCAGAGCCTGTACTGCCGCAATTGCCCTGAAGAGCCTGCAGCGGCTCCACCTGCTAAGCCCTCT
GGAGGAGAAGGGGCCCCCTGGCCTGGAAGTCAACTATGGCTCAGCTGACAACCCCCAGACCATCTG
GTTTGAGCTGCCACAGGCCCAGGAGCTGCTATACACCACTGTCTTCCTGATAGACAGCAGTGCCTC
TTGCACTGAGTGGCCCAGCATCAACTGAGAGGAGTGCAGGCCGGGGAGAGAAGAGGATGAGGCCTC
CCCCGGCCCAAGTCTACCCACATGGTCTGCCTTGATGCTATCAGATCACTGTTCTAGAACCTGC
CTCAGCACAGCCCAGCCGGCCACATGCAGGCCATGAGGCAGGGGCTGCTATCACGTACCAGCAG
GCAAAGAAAACAGCCAGACCCTCTCCAGGACGGCCTGGGGCCAAAGCGGGCTGCAGGAAGTGGCT
GGGGCACCTGAGGTTGCCAGTCTGAGGGAGATGCCACCCGACCCAGGCTCCGCCAGGCCCCA

FIG. 5B.

+

16/25

+

CATTAGCACAAGCCCAGGCATGGGAGAAACAGCTGCTGAGGAAATAAACTCCCTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

FIG. 5C.

ESTROGEN-REGULATED UNCONVENTIONAL MYOSIN-RELATED PROTEIN: COMPOSITIONS AND METHODS OF USE

+

17/25

MYQSRPGVPVPVQPSRPPKAFLRKIDPKDEALAKLINGAHSSPPMLSPSPGKGPPPAVAPRPKA
PLQLGPSSSIKEKQGPLLDLFGQKLPIAHTPPPPAPPLPLPEDPGTLAERRCLTQPVEDQGVST
QLLAPSGSVCFSYTGTPWKLFRLKEVFYPRENFSPYLLRLLCEQILRDTFSESCIRISQNERRKM
KDLLGGLEVDLDSLTTTSDSVKKRIVVAARDNWANYFSRFFPVSGESGSDVQLLAVSHRGLRLLKV
TQGPGLRPDQLKILCSYSFAEVLGVECRGGSTLELSLKSEQLVLHTARARAIEALVELFLNELKKD
SGYVIALRSYITDNCSLLSFHRGDLIKLLPVATLEPGWQFGSAGGRSGLFPADIVQPAAAPDFSFS
KEQRSGWHKGQLSNGEPGLARWDRASERPAHPWSQAHSDDSEATSLSSVAYAFLPDSHSYTMQEF
RRYFRRSQALLGQTDGGAAGKDTDSLVOYTKAPIQESLLSLSDDVSKLAVASFLALMRFMGDQSKP
RGKDEMDLLYELLKLCQOEKLRLDEIYCQVIKQVTGHPREHCTRGSFSLSLTGFFPPSTRLMPYL
TKFLQDSGPSQELARSSQEHLQRTVKYGGRRRMPPPGEMKAFLKGQAIRLLLIHLPGGVDYRTNIQ
TFTVAAEVQEELCRQMGITEPQEVQEFALFLIKEKSQLVRLPLQPAEYLN SVVVDQDVSLHSGGSTG
RPHCTSITPPTSAPTTARCCGTTFRGSCQSAPRQTRSSPGWPPCSTSARPTGI PPQGRTC

FIG. 6.

18/25

+

CGGCAGCAGCAGGCTCGGGCCTCCGAGGCTGCGTCCCAGGCTCACCTCAGCCGTCACCTCCAAG
CCCAGGAAGCCCCCACACCCCGGAGAAGCCACAGCGTGACCTGGGATCAGAGGGTGGCTGCCTG
AGGGAGACCTCCGAGGAGGCTGAAGACAGGCCCTATCAGCCCAAGAGCTTCCAGCAGAAACGGAAC
TATTTCCAGAGGATGGGGCAGCCACAGATCACAGTGAGGACGATGAAGCCCCGGCCAAGGTCCAC
ATCCCCCAGGGGGAAGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGGAGGAGCAA
GAAGTGGAACAAGAGCAGCGCCGTCCCCTCCTCCTCCCCCATCGTGAAGAAGCCATTGAAGCAA
GGTGGGGCCAAAGCTCCAAAAGAGGCTGAGGCTGAGCCAGCCAAGGAGACAGCGGCCAAGGGCCAT
GGCCAAGGGCCAGCCCAAGGCAGGGGGACTGTGGTGCGCAGTCAGACTCCAAGCCCAAGCGGCCAC
AACCAGCAGGGAAATTGGCAACATCATCCGCATGTACCAGAGCCGCCCGGGCCCCGTGCCTGTGC
CCGTGCAGCCATCCAGGCCTCCCAAAGCTTTCTGAGGAAAATCGACCCCAAGGACGAGGCTCTGG
CCAAGCTGGGTATCAACGGTGCCCACTCGTCCCCGCCGATGCTGTCCCCAGCCCAGGAAAGGGCC
CCCCGCCAGCTGTGGCTCCTCGACCCAAAGCCCCGCTACAGCTTGGGCCCTCTAGCTCCATCAAGG
AAAAGCAGGGGCCCCCTTCTGGACCTGTTTGGCCAGAAGCTGCCTATTGCCACACACCCCCACCTC
CACCAGCGCCACCACTGCCTCTGCCCCGAGGACCCAGGGACCCTTTCAGCAGAGCGTCGTTGCTTGA
CACAGCCCGTGGAGGACCAGGGGGTCTCCACCCAGCTACTCGCGCCCTCTGGCAGCGTGTGCTTCT
CCTACACCGGCACGCCCTGGAAGTTGTTTCTACGCAAGGAGGTGTTCTACCCACGGGAGAACTTCA
GCCATCCCTACTACCTGAGGCTCCTCTGTGAGCAGATCCTACGGGACACCTTCTCCGAGTCCTGTA
TCCGGATTTCCAGAATGAGCGGCGGAAAATGAAAGACCTGCTGGGAGGCTTGGAGGTGGACCTGG
ATTCTCTCACCACCACCGAAGACAGCGTCAAGAAGCGCATCGTGGTGGCCGCTCGGGACAACCTGGG
CCAATTACTTCTCCCGCTTCTTTCCTGTCTCGGGCGAGAGTGGCAGCGACGTGCAGCTGTTAGCCG
TGTCACACCGTGGGCTGCGACTGCTCAAGGTGACCCAAGGCCCCGGCCTCCGCCCCGACCAGCTGA
AGATTCTCTGCTCATAACAGCTTTGCGGAGGTGCTGGGTGTGGAGTGCCGGGGCGGCTCCACCCTGG
AGCTGTCACTGAAGAGCGAGCAGCTGGTGCTGCACACAGCCGGGCAAGGGCCATCGAGGCGCTGG
TTGAGCTATTCTGAATGAGCTTAAGAAGGACTCCGGCTATGTCATCGCCCTGCGCAGCTACATCA
CTGACAACCTGCAGCCTCCTCAGCTTCCACCGTGGGGACCTCATCAAGCTGCTGCCGGTGGCCACCC
TGGAGCCAGGCTGGCAGTTTGGCTCTGCCGGGGGCCGTTCCGGACTCTTTCCTGCCGACATAGTGC
AGCCGGCTGCCGCTCCCGACTTTTCCTTCTCCAAGGAGCAGAGGAGTGGCTGGCACAAGGGTCAGC
TGTC AACGGGGAACCAGGGCTGGCTCGGTGGGACAGGGCCTCAGAGCGCCCTGCCACCCCTTGGA
GCCAGGCACACAGTGACGACTCGGAGGCCACCAGCCTGTCTCTGTGGCCTATGCCTTTCTGCCCG
ACTCCACAGCTACACCATGCAGGAATTCGCCCGGCGTTACTTCCGGAGGTCCCAGGCCTTGCTGG
GCCAGACTGATGGAGGTGCCGAGGAAAGGACACGGACAGCCTGGTGCAGTACACCAAGGCTCCCA

FIG. 7A.

+

19/25

TCCAGGAGTCGCTCCTCAGCCTCAGTGATGATGTGAGCAAGCTGGCTGTAGCCAGCTTCCTGGCCC
TGATGCGGTTTATGGGTGACCAGTCCAAGCCCCGGGGCAAGGATGAGATGGATCTGCTCTATGAAC
TGCTGAAGCTGTGCCAGCAGGAGAAGCTGAGGGATGAGATTTACTGCCAGGTTATCAAGCAGGTCA
CAGGACACCCCCGGCCGGAACACTGCACTCGAGGCTGGAGCTTCCTCAGCCTTCTCACAGGCTTCT
TCCCCCGTCGACCAGGCTGATGCCCTACCTGACCAAGTTTCTGCAGGATTCAGGCCCCAGCCAAG
AGCTGGCCCCGAGCAGCCAGGAGCACCTCCAGCGCACAGTCAAATATGGGGGGCGCCGGCGGATGC
CCCCACCGGGTGAAATGAAGGCTTTCCTGAAAGGACAAGCGATTGCGCTGCTTCTTATTACCTGC
CGGGGGGTGTGGATTATAGGACGAATATCCAGACTTTCACAGTAGCAGCAGAAGTGCAGGAGGAGC
TGTGCCGGCAAATGGGTATCACGGAGCCTCAGGAAGTGCAGGAATTCGCCCTCTTCCTCATCAAAG
AGAAGAGCCAGCTGGTGCGGCCCCCTGCAGCCCGCCGAATACCTCAACAGCGTGGTAGTGGACCAGG
ACGTGAGCCTGCACAGCGGCGGCTCCACTGGGAGACCCCACTGCACTTCGATAACTCCACCTACAT
CAGCACCCACTACAGCCAGGTGCTGTGGGACTACCTTCAGGGGAAGCTGCCAGTCAGCGCCAAGGC
AGACGCGCAGCTCGCCAGGCTGGCCGCCCTGCAGCACCTCAGCAAGGCCAACAGGAATACCCCCTC
AGGGCAGGACCTGCTAGCTTACGTGCCAAAGCAGCTGCAACGGCAGGTGAACACGGCCTCCATCAA
GAACCTGATGGGTGAGGAGCTGAGACGGCTGGAAGGACACAGCCCCCAGGAAGCACAGATCAGCTT
CATTGAGGCCATGAGCCAGCTGCCCCCTTTCGGCTACACCGTCTATGGGGTGCTGCGAGTGAGCAT
GCAGGCCCTGTCCGACCCACTCTCCTGGGGCTCAACCGCCAGCATCTCATCCTCATGGACCCAG
CTCCCAGAGCCTGTACTGCCGCATTGCCCTGAAGAGCCTGCAGCGGCTCCACCTGCTAAGCCCTCT
GGAGGAGAAGGGGCCCCCTGGCCTGGAAGTCAACTATGGCTCAGCTGACAACCCCCAGACCATCTG
GTTTGAGCTGCCACAGGCCCAGGAGCTGCTATACACCACTGTCTTCCTGATAGACAGCAGTGCCTC
TTGCACTGAGTGGCCAGCATCAACTGAGAGGAGTGCAGGCCGGGAGAGAAGAGGATGAGGCCTC
CCCCGGCCCAAGTCTCACCCACATGGTCTGCCTTGGATGCTATCAGATCACTGTTCTAGAACCTGC
CTCAGCACAGCCCAGCCGGCCACATGCAGGCCATGAGGCAGGGGCTGCTATCACTGTCACCAGCAG
GCAAGAAAACAGCCAGACCCTCTCCAGGACGGCCTGGGGCCAAAGCGGGCTGCAGGAAGTCCGGCT
GGGGCACCTGAGGTTGCCAGTCTGAGGGAGATGCCACCCGACCCAGGCTCCGCCAGGCCCCA
CATTAGCACAAAGCCAGGCATGGGAGAAACAGCTGCTGAGGAAATAAACTCCCTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

FIG. 7B.

FIG. 8.

+

FIG. 9

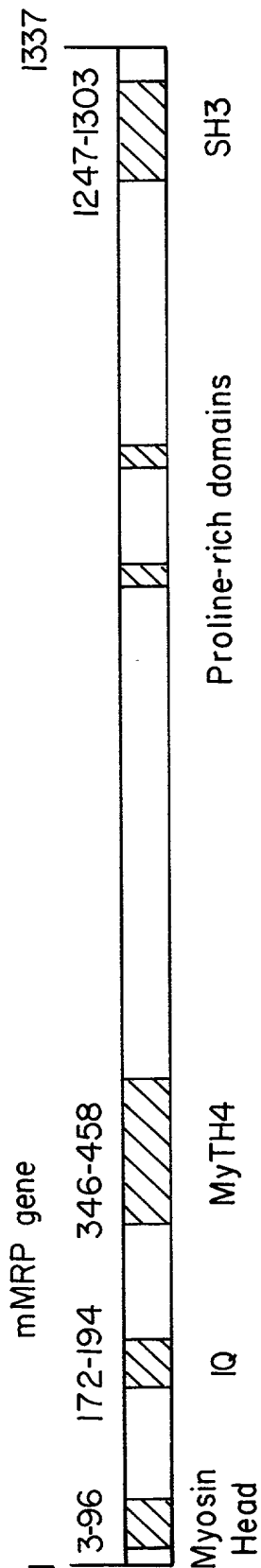


FIG. 9.

22/25

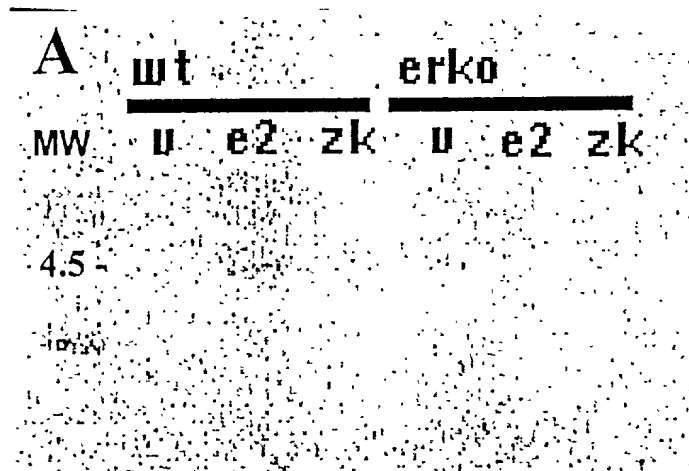


FIG. 10.

23 / 25

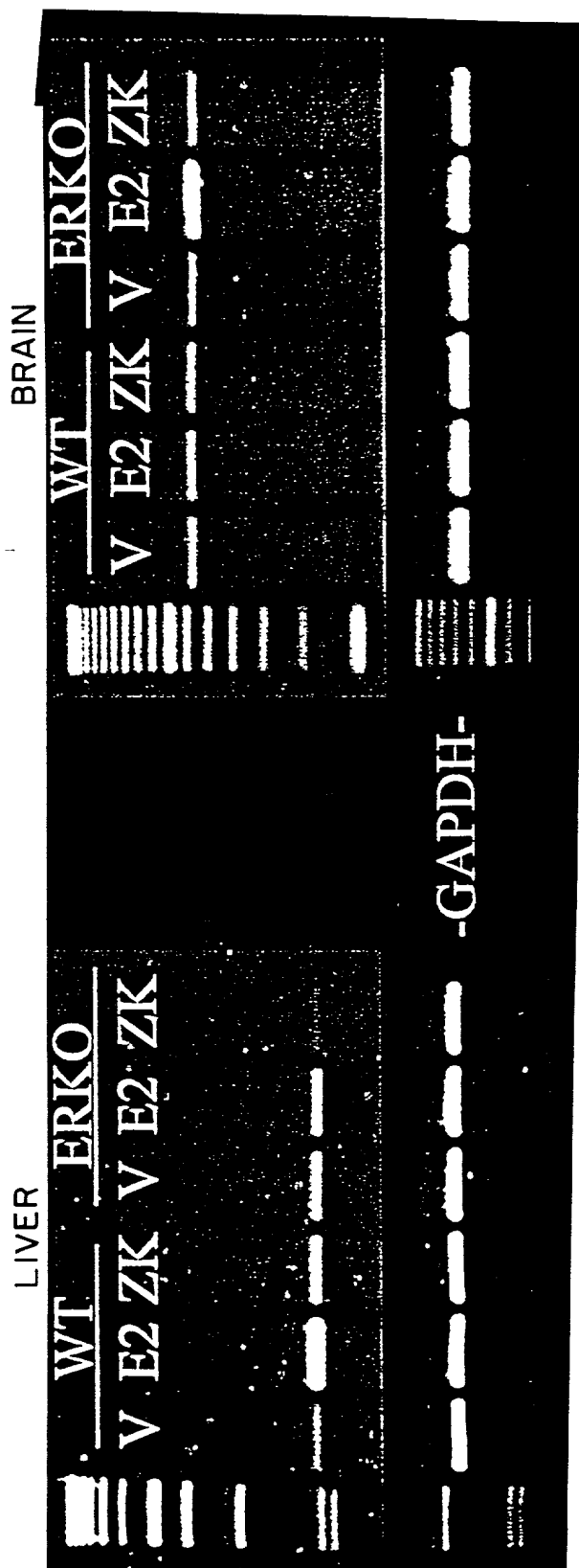


FIG. 11.

24/25

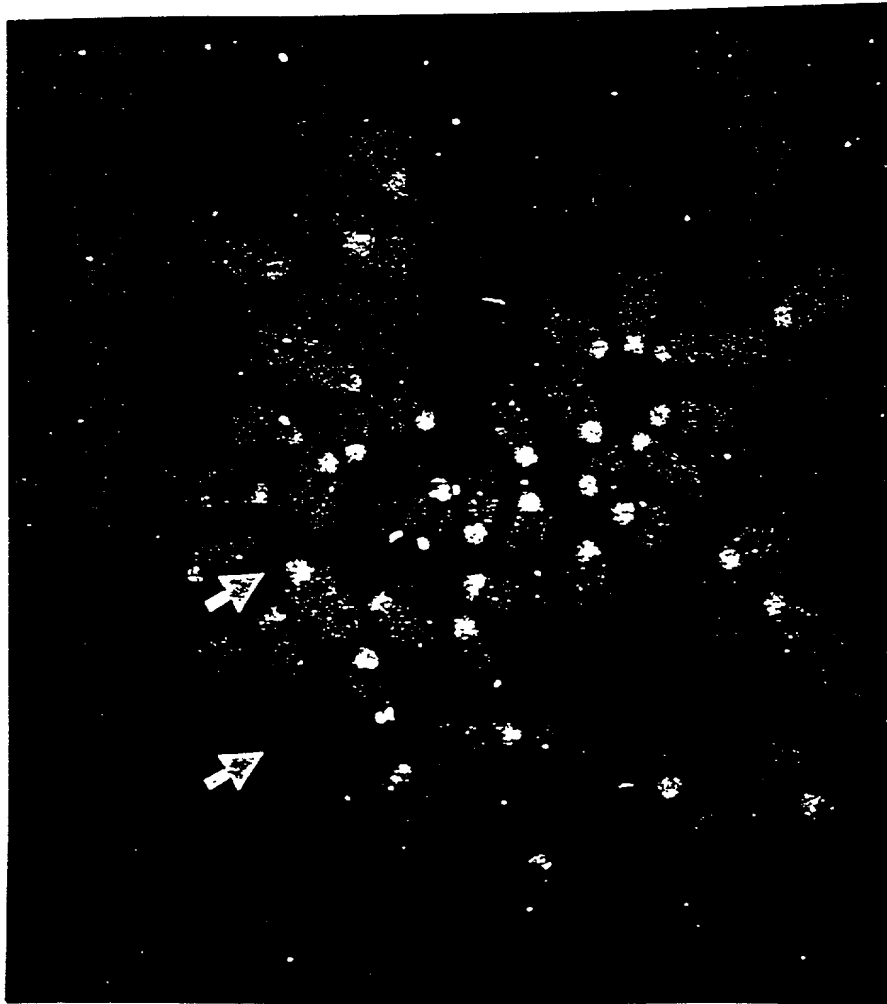
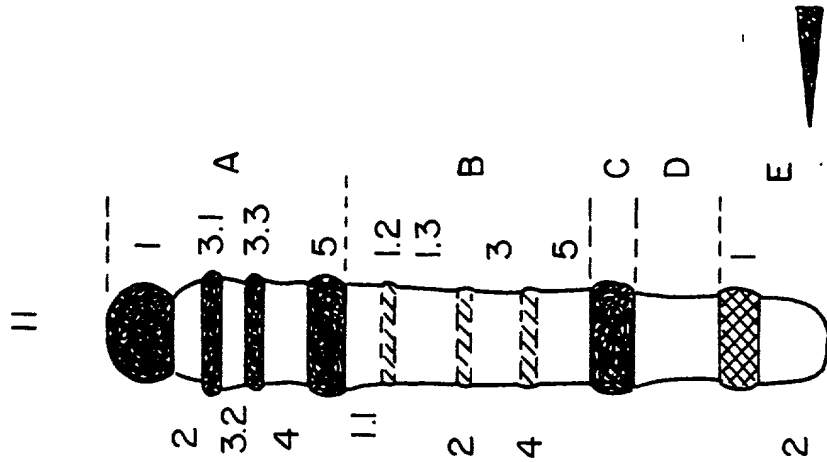


FIG. 12.

25/25

+

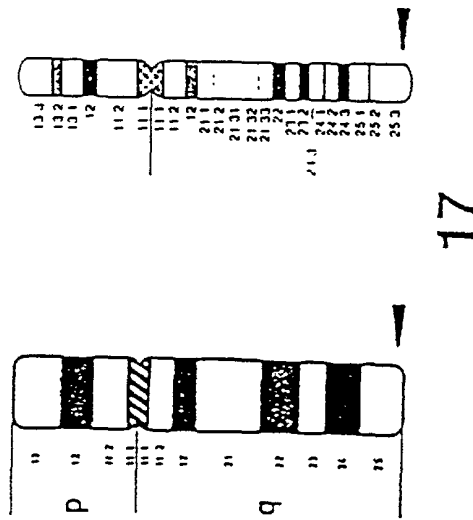


FIG. 13.